Interaction with the 3D reactive widgets for musical performance.

Florent Berthaut, Myriam Desainte-Catherine, Martin Hachet

Interacting with the 3D Reactive Widgets for Musical Performance

Florent Berthaut[1] **, Myriam Desainte-Catherine**[1] **, Martin Hachet**[2]

[1]Université de Bordeaux
Laboratoire Bordelais de Recherche en Informatique
351 cours de la libération – 33400 Talence, France

[2]INRIA
Laboratoire Bordelais de Recherche en Informatique
351 cours de la libération – 33400 Talence, France

berthaut@labri.fr, myriam@labri.fr, hachet@labri.fr

Abstract. *While virtual reality and 3D interaction provide new possibilities for musical applications, the existing immersive virtual instruments are limited to single process instruments or musical navigation tools. In this paper we present the 3D reactive widgets. These graphical elements enable simultaneous control and visualization of musical processes in 3D immersive environments. They also rely on the live-looping technique, allowing to build complex musical sequences. We describe the interaction techniques that we have designed and implemented to manipulate these widgets, including a virtual ray and tunnels. After having expressed the lack of expressivity and efficiency of the existing input devices for sound production gestures, we finally set the requirements for an appropriate device for musical interaction in 3D immersive environments.*

1. Introduction

Graphical musical interfaces have many advantages over hardware controllers. They can provide easy and direct access to a high number of parameters of an unlimited number of sound processes. They can also be used to display many useful informations about these processes. Nevertheless, they require efficient interaction devices and techniques.

The purpose of our work is to explore the possibilities provided by 3D interaction and virtual environnements. Navigation in these environments can be a good metaphor for navigation in musical pieces or exploration of musical structures. New interaction techniques developped in these fields of research open possibilities for musical interaction and can also be combined with traditional techniques. Immersion also provides new sensations to musicians and to the audience. However, these possibilities must be adapted to the specific needs of musical interaction, such as expressivity, efficiency, and minimum latency.

In this paper, we present the principle of the 3D reactive widgets. Then we describe the specific interaction techniques that we have developped. We finally set the requirements for an input device for 3D musical interaction.

2. Related Work

Relatively few research have been done in the field of 3D interaction and virtual reality for music. Some of them focus on navigation in musical environments, like the virtual groove in the Phase project [Rodet et al., 2005], or the audiovisual grains in Plumage [Jacquemin et al., 2007]. The applications developped by Mike Wozniewski et al. [Wozniewski et al., 2006] also rely on users movements to either control the spatialization of pre-recorded sound sources, or apply effects on the sound of an accoustic instrument. Some immersive instruments are single processes instruments, i.e. instruments that allow to interact with only one synthesis process, such as the Virtual Xylophone, the Virtual Membrane, or the Virtual Air Guitar developped by Mäki-Patola et al. [Mäki-Patola et al., 2005] and the sculpting instruments developped by Mulder [Mulder, 1998]. Finally, among the existing multi-processes 3D instruments, part of them, like the WAVE software from Valbom et al. [Valbom and Marcos, 2005] or the application developped by Martin Naef et al. [Naef and Collicot, 2006], have limited visual feedback and interaction possibilities since they tend to emulate hardware controllers. The other part of these instruments rely on gaming software or devices, like the 3D instruments Fijuu [Olive and Pickles, a] and Ergates [McCormick,], or the collaborative musical First Person Shooters q3apd [Olive and Pickles, b] and q3osc [Hamilton, 2008]. They offer new interaction techniques and interesting visualizations, but they do not take advantage of the possibilities of immersive environments. None of these applications combine immersion, simultaneous control of multiple processes, expressive interaction techniques and complex visual feedback.

3. The 3D Reactive Widgets

3.1. Principle

Our research focus on using 3D immersive environments for musical interaction. These environments indeed add possibilities in terms of temporal or hierarchical navigation by means of 3D movements. They also enable the design of new interaction techniques and paradigms, for example using the additional dimensions for manipulation and visualization. Furthermore, immersion can improve the experience of the musicians who will better perceive and thus manipulate the 3D interface with stereoscopic display and head-tracking. But it will also improve the experience of the audience, if they are equipped with stereoscopic glasses, both for the spectacular aspect and for the understanding of the musicians playing.

As said in the previous section, most virtual reality instruments are single process instruments, i.e instruments that allow the control of only one synthesis or effect process, or musical navigation tools. However, we believe that the main advantage of graphical musical interfaces is to give the possibility to handle multi-processes instruments with control on and visual feedback from the selected sound processes.

This is why we chose to rely on the concept of reactive widgets described by Golan Levin[Levin, 2000] and used for example by Sergi Jorda[Jordà, 2005] in FMOL. A reactive widget is a graphical component which allows both manipulation and visualization of a musical process. Its graphical parameters are connected to the parameters of the associated musical process. These connections are bidirectionnal, so that graphical changes are reflected in the sound process and that musical events are displayed in return by the widget. The efficiency of this concept lies in the shortening of the "indirection degree" described by Michel Beaudoin-Lafon[Beaudoin-Lafon, 1999] because there is direct manipulation of the "objects of interest", in our case the visualized sound processes.

These observations led us to adapt the concept of the reactive widgets to 3D immersive environments, as it can be seen in figure 1.

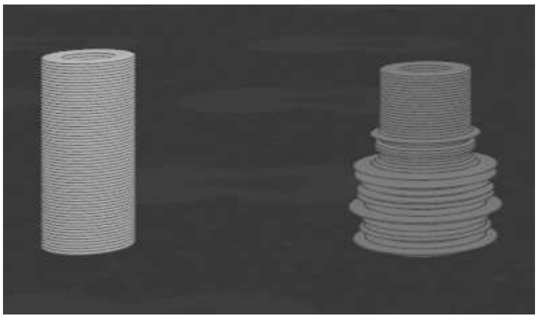

Figure 1: Two 3D Reactive Widgets: the one on the right shows the spectrum of its associated sound process.

3.2. Audiovisual Mappings

Several questions emerge from the concept of the 3D reactive widgets. First of all, one must choose which sound and musical parameters should be controlled and visualized. Symetrically, one must choose which graphical parameters should be used to manipulate and display the sound processes. Finally, one must choose the mappings between these audio and graphical parameters. These questions have been explored in several user studies, such as [Giannakis, 2006], but always in terms of users preferences and with a music composition perspective. This is why we are currently running a specific user study aimed at interaction, which will focus on mappings efficiency by measuring subjects' performances. Meanwhile, the following parameters and mappings were chosen for our first implementation: spectrum with shape, color lightness with pitch, size with loudness, shape distortion with brightness, and transparency with noisiness.

3.3. Sound processes

The sound process attached to each 3D reactive widget is composed of a sample player, an audio effects rack and an audio analysis rack. The widget, when activated by a specific manipulation, triggers the sample. The sound goes through the effects rack, whose parameters are linked to the graphical parameters of the widget according to the mappings described in the previous subsection. The result is then sent to the soundcard output, and at the same time analyzed to set the widget's shape. A very important feature of our application is that the sound triggers and the manipulations of the graphical parameters can be recorded and looped with the live-looping technique. This technique is used by a growing number of musicians in every musical genre and relies on music software or hardware devices. In our case, when the musician grabs a widget, he or she can press a button to start recording. All the following parameters variations are recorded, without temporal quantization, and looped when the button is pressed again. All the loops of a widget are synchronized, and can easily be deleted. This allows to build, stack and manipulate complex sound processes using the 3D reactive widgets.

4. Implementation

In the current implementation, as it can be seen in figure 3, the musician is equipped with head-tracked stereoscopic glasses and uses tracked Wii Remotes. The display is a large screen combined with Infitec stereo projectors and the tracking is done with the A.R. Tracking 6DOF DTrack system[1]. A cheaper system could be set up with simple anaglyph stereoscopy combined with tracking using a wiimote IR sensor. Tracking data is transmitted using VRPN[2].

[1] http://www.ar-tracking.de/

[2] http://www.cs.unc.edu/Research/vrpn/

The 3D environment is rendered by an application which we have developped, called Poulpe. Poulpe is based upon the OpenSG scenegraph library[3] and the oscPack library[4]. It can be seen as a bridge between virtual reality and music software because it associates musical control messages with each parameter of each element of a 3D scene. Thus the camera, the lights and the 3D objects can be used to control and visualize musical processes. The scene is described in a text file with an xml syntax.

In our application, Poulpe communicates with custom music software using the OpenSoundControl[5] protocol. The music application uses the Jack sound server[6]. Sound files are associated with each 3D reactive widget. Effects are applied using LV2 plugins[7]. Finally, audio analysis is done with Vamp plugins [8]. The use of plugins allows to easily add new effects and sound features. The resulting setup is shown in figure 2.

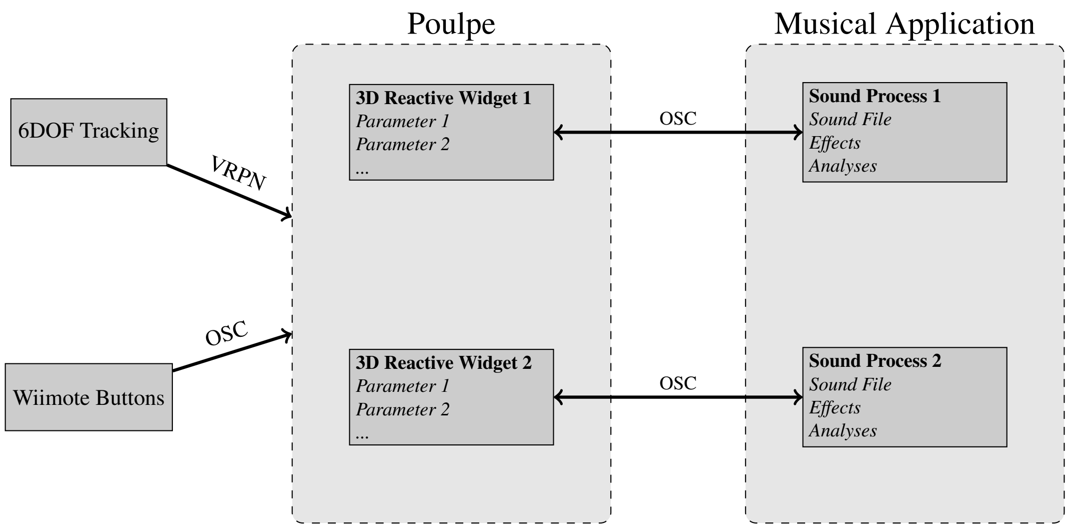

Figure 2: Block diagram of the current implementation.

5. Interacting with the 3D reactive widgets

In the next subsections, we describe the first interaction techniques chosen to manipulate the 3D reactive widgets. These techniques can be seen in figure 3 and in a video[9]. They can be categorized, by refering to Cadoz's [Cadoz, 1999] work on musical gestures, as selection gestures, modulation gestures and excitation gestures.

5.1. Virtual Ray

In order to select and to grab the reactive widgets, the musician manipulates a virtual ray, which is commonly used in virtual reality applications [Bowman et al., 1997]. It was evaluated as an efficient technique especially for near objects, by Poupyrev et al. [Poupyrev et al., 1998]. It gives a feeling of continuity from the real world and provides sufficiently accurate and fast pointing in the virtual environment. In our application, the movements of the ray are low-pass filtered when over a non-grabbed reactive widget, to

[3] http://opensg.vrsource.org/trac
[4] http://www.audiomulch.com/ rossb/code/oscpack/
[5] http://opensoundcontrol.org/
[6] http://jackaudio.org/
[7] http://lv2plug.in/
[8] http://www.vamp-plugins.org/
[9] http://rapidshare.com/files/249367217/interacting-3D-reactive-widgets-musical-performance.mpg.html

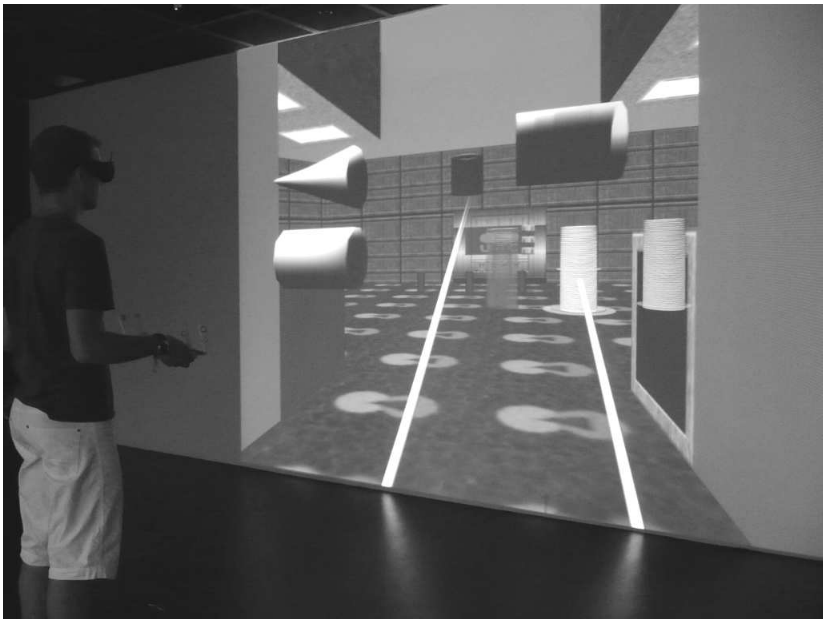

Figure 3: The current application with four 3D reactive widgets, three tunnels and two virtual rays.

avoid unwanted jumps while triggering the widget. We also chose to limit the translations applied to the widget to the X and Y axes to make the selection and manipulation gestures easier keeping the movements on the Z axis for navigation in the 3D environment. Two rays can finally be used at the same time, one in each hand. This technique corresponds to the selection gestures, by which musicians choose the instruments components, like the key on a piano.

5.2. Tunnels

In order to manipulate the widgets' graphical parameters, the musician moves them through what we call "tunnels". Each tunnel modifies one graphical parameter, and thus the corresponding sound parameter, following specific scales. Scales presets, such as non-linear or discretized scales, can be defined in the Poulpe configuration file and selected while playing. Our current application is thus composed of a Size Tunnel, a Distortion Tunnel, a Transparency Tunnel and a Color Tunnel. They can be seen in figure 4. As indicated in the previous section, these variations can be live-looped. This technique corresponds to the parametric modulation gestures, which modify the properties of an instrument, such as the note on a string of a violin.

5.3. Sounds Trigger

Finally, in order to trigger the sounds of the 3D reactive widgets, two techniques were experimented: the buttons of the wii remote and the collisions of the widgets (with vibro-tactile feedback). The buttons allow for simultaneous excitation and modulation gestures, but they lack expressivity and temporal accuracy. Triggering the sounds by hitting the widgets against other objects is more expressive, because the speed of the collision can be used. But, on the other hand, it increases the latency and requires more accurate active haptic feedback. These two techniques are thus not efficient enough to be used as the excitation gestures, i.e. gestures that directly and physically generate the sound. This was indeed confirmed during the demonstrations that we made for the VRST 2008 conference. In order to enable efficient musical interaction in immersive environments, a new device must be designed.

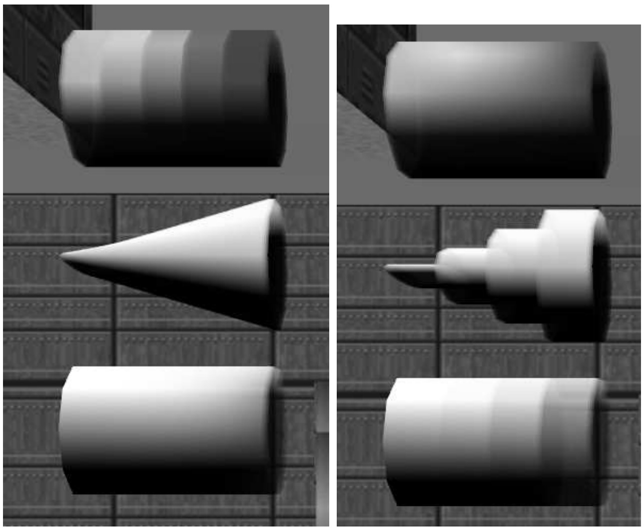

Figure 4: Three Tunnels: Color/Pitch, Size/Volume, Transparency/Noisiness. By sliding a 3D reactive widget through them, the musician modifies its graphical parameters, and thus the corresponding parameters of the associated sound process. The second figure shows a different preset for each tunnel.

6. Towards a new interaction device

6.1. Related devices and limitations

The interaction device currently used in our application is a wii remote equipped with 6DOF targets. This device, as other videogames controllers such as joysticks and gamepads, does not provide controls that are expressive and accurate enough. Moreover, the bluetooth protocol used for data transmission adds an average of 35ms of latency. Virtual reality devices, such as the flystick from A.R Tracking, focus on accurate orientation and position sensing and does not provide expressive controls. Advanced hardware instruments, for example the ones presented in the NIME conferences[10], usually focus on expressivity. However they are not usable for graphical interaction because either they are designed for single processes instruments, or they are not generic enough, or they are not handheld and limit the musicians movements. The Meta-Instrument[11] could be a good solution, since it provides a lot of expressive inputs and sensing of arms movements. However, it does not give absolute 3D position and orientation of the musician and the control of the virtual ray with the forearm would be far less accurate and fast than with the wrist. Advanced haptic controllers such as the one designed by the ACROE[12] restrict the movements in translation and rotation and thus are not desirable for interaction in immersive environments, especially for the control of a virtual ray. Finally, the few devices specifically designed for immersive musical interfaces like the device for Ashitaka [Moody, 2006] or the Sphere Spatializer [Wakefield et al., 2008] focus on navigation and do not enable expressive excitation gestures.

6.2. Requirements

Considering the limitations of existing devices, there is a need for an efficient device for musical interaction in immersive environments. We believe that it should meet the following requirements:

[10]http://www.nime.org/
[11]http://www.pucemuse.com
[12]http://acroe.imag.fr/ergos-technologies/index.php?idp=0

Figure 5: Using wii remotes with 6DOF tracking to control the virtual rays

- It should provide accurate and fast 6DOF tracking.
- It should allow simultaneous excitation and modulation gestures.
- The communication protocol used should minimize latency, so that percussive gestures can be correctly achieved.
- The number of controls available and their expressivity should be maximized.
- It should be easily handheld.
- It should be sufficiently generic to enable the design of new interaction techniques and should be usable for non-musical applications.
- Finally, it should provide haptic feedback (at least passively).

7. Conclusion

We described a new concept for graphical musical interaction: the 3D reactive widgets. Each of these widgets allow both the control and the visualization of a musical process. This principle is combined with the possibilities provided by 3D immersive environments in terms of interaction and immersion. It is also combined with the live-looping technique, allowing the creation of complex musical sequences. We introduced the first interaction techniques designed to manipulate these widgets, which rely on the virtual ray for the selection gestures, and on what we have called tunnels for modulation gestures. In the last section, we expressed the lack of an appropriate device for the excitation gestures, i.e. gestures that generate the sound. Further work will be to design this new interaction device and new techniques that it will enable. This device could also be used for other immersive applications, even non-musical ones that requires accuracy and expressivity. Finally, our in-progress user study on audiovisual mappings may help choosing the right mappings for various graphical musical interfaces, including our application.

References

Beaudoin-Lafon, M. (1999). *Interfaces homme-machine et création musicale*, chapter Moins d'interface pour plus d'interaction. Hermès Sciences.

Bowman, D., Koller, D., and Hodges, L. F. (1997). Travel in immersive virtual environments: An evaluation of viewpoint motion control techniques. In *IEEE Proceedings of VRAIS97*.

Cadoz, C. (1999). *Musique, geste, technologie*. Éditions Parenthèses.

Giannakis, K. (2006). A comparative evaluation of auditory-visual mappings for sound visualisation. *Organised Sound*.

Hamilton, R. (2008). q3osc: or how i learned to stop worrying and love the game. In *Proceedings of the International Computer Music Association Conference*.

Jacquemin, C., Ajaj, R., Cahen, R., Ollivier, Y., and Schwarz, D. (2007). Plumage: Design d'une interface 3d pour le parcours d'échantillons sonores granularisés. In *Proceedings of the Conférence Francophone sur l'Interaction Homme-Machine(IHM'07)*.

Jordà, S. (2005). *Crafting musical computers for new musics' performance and improvisation*. PhD thesis, Universitat Pompeu Fabra.

Levin, G. (2000). *Painterly Interfaces for Audiovisual Performance*. PhD thesis.

Mäki-Patola, T., Laitinen, J., Kanerva, A., and Takala, T. (2005). Experiments with virtual reality instruments. In *Proceedings of the 2005 International Conference on New Interfaces for Musical Expression (NIME05), Vancouver, BC, Canada*.

McCormick, C. Ergates: http://mccormick.cx/projects/ergates/.

Moody, N. (2006). Motion as the connection between audio and visuals.

Mulder, A. G. (1998). *Design of virtual three-dimensional instruments for sound control*. PhD thesis.

Naef, M. and Collicot, D. (2006). A vr interface for collaborative 3d audio performance. In *Proceedings of the 2006 International Conference on New Interfaces for Musical Expression (NIME06), Paris, France*.

Olive, J. and Pickles, S. Fijuu: http://www.fijuu.com/.

Olive, J. and Pickles, S. q3apd: http://selectparks.net/ julian/q3apd/.

Poupyrev, I., Weghorst, S., Billinghurst, M., and Ichikawa, T. (1998). *Egocentric object manipulation in virtual environments: Empirical evaluation of interaction techniques*.

Rodet, X., Gosselin, F., Mobuchon, P., Lambert, J.-P., Cahen, R., Gaudy, T., and Guedy, F. (2005). Study of haptic and visual interaction for sound and music control in the phase project. In *Proceedings of the 2005 International Conference on New Interfaces for Musical Expression (NIME05), Vancouver, BC, Canada*.

Valbom, L. and Marcos, A. (2005). Wave: Sound and music in an immersive environment. *Computer and Graphics*.

Wakefield, G., Morin, J. K., Novak, M., Overholt, D., Putnam, L., Thompson, J., and Smith, W. (2008). The allobrain: an interactive stereographic, 3d audio immersive environment. In *CHI Conference Workshop on Sonic Interaction Design*.

Wozniewski, M., Settel, Z., and Cooperstock, J. (2006). A spatial interface for audio and music production. In *Proceedings of the International Conference on Digital Audio Effects (DAFx), 2006*.